# AD-A205 490

(2)

## REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

| | |
|---|---|
| 1a. REPORT SECURITY CLASSIFICATION<br>Unclassified | 1b. RESTRICTIVE MARKINGS<br>DTIC FILE COPY |
| 2a. SECURITY CLASSIFICATION AUTHORITY | 3. DISTRIBUTION/AVAILABILITY OF REPORT<br>Distribution Unlimited |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S) | 5. MONITORING ORGANIZATION REPORT NUMBER(S) |
| 6a. NAME OF PERFORMING ORGANIZATION<br>Air Force Operational Test & Evaluation Center (AFOTEC)<br>6b. OFFICE SYMBOL (If applicable)<br>HQ AFOTEC | 7a. NAME OF MONITORING ORGANIZATION |
| 6c. ADDRESS (City, State, and ZIP Code)<br>Kirtland AFB NM 87117-7001 | 7b. ADDRESS (City, State, and ZIP Code) |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION<br>Same<br>8b. OFFICE SYMBOL (If applicable)<br>Same | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER |
| 8c. ADDRESS (City, State, and ZIP Code)<br>Same | 10. SOURCE OF FUNDING NUMBERS<br>PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO. |

11. TITLE (Include Security Classification)
AFOTEC Technical Journal, Volume 2, Number 2, July 1988

12. PERSONAL AUTHOR(S)

| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
|---|---|---|---|
| Journal | FROM ____ TO ____ | 880700 | 32 |

16. SUPPLEMENTARY NOTATION
Periodical (2 issues/yr)

| 17. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | Test and Evaluation |
| 14 | 02 | | Operations Research |

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

DTIC COPY INSPECTED 1

DTIC ELECTE 28 FEB 1989 D E

| Accession For | |
|---|---|
| NTIS GRA&I | X |
| DTIC TAB | ☐ |
| Unannounced | ☐ |
| Justification | |
| By | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A-1 | |

| | |
|---|---|
| 20. DISTRIBUTION/AVAILABILITY OF ABSTRACT<br>☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS | 21. ABSTRACT SECURITY CLASSIFICATION<br>Unclassified |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL<br>K R Newsom | 22b. TELEPHONE (Include Area Code)<br>(505)846-2574 (AV246-)<br>22c. OFFICE SYMBOL<br>HQ AFOTEC/RS |

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

# AIR FORCE OPERATIONAL TEST AND EVALUATION CENTER

VOLUME 2 NUMBER 2 JULY 1988

## Table of Contents



Keywords:


# AFOTEC Technical Journal

The AFOTEC Technical Journal is a funded Department of the Air Force field command periodical published twice a year for more than 1,000 people assigned or attached to the Air Force Operational Test and Evaluation Center at Kirtland AFB, N.M., 87117-7001, at a three people per copy ratio. Opinions expressed herein do not necessarily represent those of the United States Air Force.

Commander .................... Maj. Gen. Cecil W. Powell
Vice Commander .................. Col. Joseph E. Merrick
Director of Public Affairs....... Capt. Garrett T. Mason
Editor .......................... Dr. Marion L. Williams
Assistant Editor ................ SSgt. Truevenia Carr

# Thoughts from the Chief Scientist

With the third issue of the Technical Journal, you should see a marked improvement in format. It has changed from a newsletter to a professional journal, with more improvements forthcoming.

The purpose of the journal is to disseminate technical information concerning test and evaluation, and to discuss policy issues that are technically oriented. Conceptually, OT&E is a simple mission; put the system in a realistic environment with operational personnel and determine whether it is effective and suitable. In other words, see if it can do its mission in the expected environment when operated and maintained as planned. But simple in implementation OT&E is not. We can seldom replicate the mission scenario, and must deal with the normal uncertainty and non-repeatability of field testing, problems of immaturity, less-than-representative configurations, etc. Certainly, OT&E is a complex problem, but a tractable one.

There are several ways to approach OT&E. We can use whatever we can find as a representative environment, and then put the system in that environment and see what happens. At the other end of the spectrum, we can create a true operational environment, then use a statistically valid full factorial test design with all controlled and uncontrolled variables completely instrumented. We shouldn't do the former, and probably can't afford the latter, but in no case is gut feel operational testing acceptable.

The secret of effective OT&E is to develop a well defined and comprehensive test concept, far in advance of writing the test plan. The concept should contain not only the issues and objectives, but also the basic methodology, test scenarios, how (if) simulation will be used to complement the test and required test resources. Embedded in the test concept should be a test design detailing the key system performance factors (speed, altitude, day night, etc.), sample size, number of test replications, etc. Integral to the design must be a clear understanding of why these factors are important.

For example, suppose we choose a sample size of one--a perfectly valid partial factorial design. We should be able to support that choice with sound rationale which includes an analysis of its impact on drawing valid test conclusions. The rationale may well be that each scenario can only be repeated once. As another example, consider a system for which the user performance criteria requires that target locations be resolved to within .25 nm. Only by fully understanding instrumentation errors, estimated variance, and planned analysis techniques will we be able to identify test design strengths and weaknesses to yield well-grounded conclusions on system effectiveness.

After weighing such test design factors, we may be justified to request more test time and/or resources. At the very least, we will know what to expect. The topic of more structured testing is not merely academic; to paraphrase General Powell, AFOTEC must accomplish more detailed, thorough test planning and it must be done earlier in the life of a test program.

Future issues of the Technical Journal will address issues such as test design, analysis techniques and other related topics. If you are familiar with specific examples of good test design or even test design problems, please share them with us. Sound test design is critical to our business-- effective OT&E.

# A Vector-Normalization in TOPSIS That Gives a Rank Ordering Which is Invariant Under a Special Class of Attributes

By Srikantan S. Nair, Aviation Division,
Operational Test and Evaluation Agency,
Falls Church, Va.

## INTRODUCTION

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is a method used in Multiple Attribute Decision Making (MADAM) (1). It rank orders a set of alternatives based on a set of attributes and their associated weights. TOPSIS is based on the principle that the alternative closest to the ideal solution, while farthest from the negative ideal solution, is the preferred one. The ideal solution consists of the "best" attribute values available irrespective of the alternatives. In the TOPSIS method described in (1), the decision matrix (or the matrix of attribute values) is normalized to make all attributes the same unit length vector as well as independent of the units measurement.

## BASIC TOPSIS

A summarization of the TOPSIS method described in (1) follows. If $A_i$ is the i-th alternative and $C_j$ the j-th attribute and $x_{ij}$ the measure (quantitative) of the j-th attribute for the i-th alternative, then the decision matrix is defined by $(x_{ij})$, i = 1 to m and j = 1 to n; m is the number of alternatives and n is the number of attributes. As a first step of TOPSIS, each attribute is normalized to make all attributes the same unit length of vector as well as independent of the units of measurement. Thus the normalized decision matrix is given by

$$\left( x_{ij} \Big/ \sqrt{\sum_{i=1}^{m} x_{ij}^2} \right) \qquad (1)$$

In the second step, the measures of each attribute in the normalized decision matrix are multiplied by its weight to obtain a weighted normalized decision matrix

$$(v_{ij}) = \left( w_j * x_{ij} \Big/ \sqrt{\sum_{i=1}^{m} x_{ij}^2} \right), \qquad (2)$$

i = 1 to m and j = 1 to n,

where $w_j$ is the weight of the j-th attribute. Computational routines will hold good even if the weights do not sum to 1 since when the final ratio $q_i$ is computed in equation (5), the required normalization factor will be cancelled out. Of course, to maintain credibility of the TOPSIS weights, it should be standard practice to have the weights sum to 1. The weighted normalized numbers $v_{ij}$ are used to find the set of best and the worst attribute values.

Attributes are classified into one of two types: benefit attributes and cost attributes. If the measure of an attribute is such that larger is better, then it is classified as a benefit attribute. If the measure is such that smaller is better, then it is classified as cost attribute. The best (e.g., ideal) attainable measure of a benefit attribute $C_j$ is $v_j^+ = \max_i v_{ij}$, and its worst (e.g., negative-ideal) attainable measure is $v_j^- = \min v_{ij}$. On the other hand, the best attainable measure of a cost attribute $C_k$ is $v_k^+ = \min v_{ik}$, and its worst attainable measure is $v_k^- = \max v_{ik}$.

In the third step, the separation measures of each alternative from the ideal measure,

$$S_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2}, \qquad (3)$$

i = 1,2,...,m

$$S_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^-)^2} \qquad (4)$$

i = 1, 2 ,..., m

are computed. Finally, we calculate the relative closeness of alternative $A_i$ from the ideal measure

$$q_i = S_i^- / (S_i^+ + S_i^-)$$

$$i = 1, 2, \ldots, m \qquad (5)$$

where $q_i$ takes a value between 0 and 1. Also, $q_i$ equals 1 when alternative $A_i$ has the best measure of every attribute and 0 when $A_i$ has all the worst measures. Of course, it is a rare event to have $q_i = 1$ or 0.

## PROBLEM

There is a special class of attributes which can be converted to cost (or benefit) from benefit (or cost) by subtracting its measure from a constant. For example, if an attribute is measured in percent, by subtracting its measure from 100, the attribute will be changed from benefit to cost (or cost to benefit). Similarly, the attributes measured as probabilities when subtracted from 1, or attributes measured as counts when subtracted from total counts where total counts are the same for all alternatives, change the cost-benefit status. TOPSIS rank orders should not depend on how this special class of attributes are recorded, as cost or benefit. The recording of these attributes as cost or benefit must be a choice of convenience. In this paper we propose a vector normalization which gives a TOPSIS rank ordering that is independent of whether the special class of attributes are recorded as cost or benefit.

## MODIFIED TOPSIS METHODOLOGY

In place of the normalization (1), we propose the following normalization

$$(x_{ij} - \bar{x}_j) \Big/ \sqrt{\sum_{i=1}^{m} (x_{ij} - \bar{x}_j)^2} \qquad (6)$$

which gives a weighted normalized decision matrix (7) in place of (2).

$$(\tilde{v}_{ij}) = \left( w_j * (x_{ij} - \bar{x}_j) \Big/ \sqrt{\sum_{i=1}^{m} (x_{ij} - \bar{x}_j)^2} \right), \qquad (7)$$

where $w_j$ is the weight and $\bar{x}_j$ the mean of the j-th attribute. If the j-th attribute is benefit then $\tilde{v}_j^+ = \max \tilde{v}_{ij}$ is its best attainable measure and $\tilde{v}_j^- = \min \tilde{v}_{ij}$ is its worst attainable measure. On the other hand, if the j-th attribute is cost, then $\tilde{v}_j^+ = \min \tilde{v}_{ij}$ is its ideal and $\tilde{v}_j^- = \max \tilde{v}_{ij}$ is its negative-ideal attainable measure.

If $\tilde{s}_i^+$ denotes the separation measure of the i-th alternative from the ideal measure and $\tilde{s}_i^-$ that from the negative ideal, then

$$\tilde{s}_i^+ = \sqrt{\sum_{j=1}^{n} (\tilde{v}_{ij} - \tilde{v}_j^+)^2} \qquad (8)$$

$$i = 1, 2, \ldots, m$$

and

$$\tilde{s}_i^- = \sqrt{\sum_{j=1}^{n} (\tilde{v}_{ij} - \tilde{v}_j^-)^2} \qquad (9)$$

$$i = 1, 2, \ldots, m$$

Finally the relative closeness of the i-th alternative from the ideal measure is given by $\tilde{q}_i$, where

$$\tilde{q}_i = \tilde{s}_i^- / (\tilde{s}_i^+ + \tilde{s}_i^-) \qquad (10)$$

$$i = 1, 2, \ldots, m.$$

The set of alternatives can now be rank ordered using these $\tilde{q}$-values. The alternative which has the highest $\tilde{q}$-value is the most preferred and the one with lowest $\tilde{q}$-value is the least preferred. The rank ordering based on $\tilde{q}$-values will be the same regardless of how the special class of attributes are recorded, cost or benefit.

## NUMERICAL EXAMPLE

The decision matrix of the navigation data on four aircraft is given below. The alternatives are the four aircraft A, B, C, and D and the attributes are the six measures of performance (MOP).

$(x_{ij}) =$ (11)

| MOP1 | MOP2 | MOP3 | MOP4 | MOP5 | MOP6 | ALTERNATIVE |
|---|---|---|---|---|---|---|
| 100.0 | 100.0 | 71.7 | 71.1 | 87.1 | 72.9 | A |
| 90.0 | 50.0 | 63.8 | 41.3 | 92.3 | 54.9 | B |
| 100.0 | 50.0 | 40.1 | 36.0 | 63.6 | 76.7 | C |
| 100.0 | 100.0 | 55.0 | 104.5 | 76.2 | 75.7 | D |

where MOP4 is cost and all others are benefit attributes. The weights of the MOPs are given by the weight vector:

$$w = (.1979,.1505,.1657,.1275,.2079,.1468)$$

In the above decision matrix, the attributes are defined as follows.
MOP1: Percent courses completed (benefit); MOP2: Percent courses not disoriented (benefit); MOP3: Percent of time within 200m of course (benefit); MOP4: Self-location error in meters (cost); MOP5: Percent self-locations within 200m (benefit); and MOP6: Velocity in knots (benefit). The attributes MOP1, MOP2, MOP3, and MOP5 belong to the Special Class defined earlier, since each is a percent.

The rank order obtained using the basic TOPSIS normalization is A, B, C, and D where alternative A is most preferred and D is least preferred. Their q-values were .71, .58, .48, and .46, respectively. When the decision maker (e.g., evaluator) reclassified MOP2 and MOP3 as cost attributes along with MOP4 and all others remained benefits, the decision matrix became

$$(x_{ij}) = \begin{array}{cccccccc} MOP1 & MOP2 & MOP3 & MOP4 & MOP5 & MOP6 & & \\ 100.0 & 0.0 & 29.3 & 71.1 & 87.1 & 72.9 & A & \\ 90.0 & 50.0 & 36.2 & 41.3 & 92.3 & 54.9 & B & (12) \\ 100.0 & 50.0 & 59.9 & 36.0 & 63.6 & 76.7 & C & \\ 100.0 & 0.0 & 45.0 & 104.5 & 76.2 & 75.7 & D & \end{array}$$

The weight vector is same:

$$w = (.1979,.1505,.1657,.1275,.2079,.1468)$$

MOP2 is now percent of courses disoriented (cost) and MOP3 is percent of time <u>outside</u> 200m of course (cost).

The rank order now obtained in the case (still using the basic TOPSIS normalization) was A, D, B, C, with q-values .80, .61, .43, and .35, respectively. Thus, D moved to second place from the fourth place when MOP2 and MOP3 were changed from benefit to cost.

In the case of modified TOPSIS normalization technique described in this paper, both the decision matrices (11) and (12) gave the same rank order: A, D, B, C. The $\tilde{q}$-values for A, B, C, and D in both cases were .82, .51, .47, and .59, respectively. It is recommended that the modified vector - normalization presented in this paper be used whenever one or more of the attributes in the TOPSIS analysis belong to the special class of attributes.

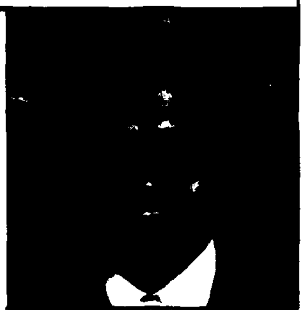
## SOFTWARE

An in-house software package has been developed to compute the q and $\tilde{q}$ values which can be run on an IBM PC or IBM compatible PC. There is no limitation on the number of alternatives and the number of attributes. Copies of the software can be obtained by calling AV 289-2464/96.

## REFERENCE

(1) Hwang, C. and Yoon, K., "Multiple Attribute Decision-Making Methods and Applications," <u>Lecture Notes in Economics and Mathematical Systems.</u> Springer-Verlag, New York 1981, pp. 128-140.

**Dr. Srikantan S. Nair** is the Mathematical Statistician, Aviation Division, U.S Army Operational Test and Evaluation Agency. His major duties involve applying mathematical and statistical methodologies in operational test plans, and developing methods and techniques through which system performance and operational interrelationships can be valued and measure. He has published several papers on operations research. Dr. Nair holds a bachelor's in Mathematics from University of Kerala, India and a master's and Ph.D. in Mathematical Statistics from Purdue University.



# The Analytical Test and Evaluation Software System (ATESS)

By Ms. Pat L. Brannen,
Operations Analysis Directorate and
Ms. Kathleen M. Hibson,
ATESS Program Manager BDM Corp.

## INTRODUCTION

ATESS was developed for AFOTEC to meet a wide range of typical T&E data reduction and analysis requirements. The need for something like ATESS became apparent to analysts and software developers through experience gained supporting a number of test programs, such as Wild Weasel (WW), Precision Location Strike System (PLSS), and High-Speed Anti-Radiation Missile (HARM).

Each test program began with the same challenge: develop data reduction and analysis support software with little or no documentation of the "real" data structures and relationships that would be available during the test. This uncertainty often delayed the start of the software development process until it was too late to get done before the test started.

Even when the software development process began early enough to ensure adequate time and resources, the test program could still be delayed while software was revised to accommodate last minute changes to poorly documented data formats and relationships.

The same types of software requirements, with minor variations, were being identified and addressed on different test programs. Each new program's requirements were solved by developing a new software package tailored to that single program. Many of the processes were common to several test programs; but specific data structure differences for each program required different software solutions, preventing use of existing data reduction and analysis software to support new test programs or requiring extensive revisions to meet new constraints.

## CONCEPTUAL DESIGN

ATESS is a comprehensive set of modular tools, designed to provide flexible data reduction and analysis support for field tests. ATESS can be adapted to changing data formats and flows without changing the code itself. Rather, changes are made through use of external files containing such information as record type, length, and format; delimiters; numbers and types of parameters; valid data ranges; functions to be performed; and relationships between data sets and contexts. In the past, these types of information were usually embedded within typical data processing and analysis software code, requiring time-consuming software revisions each time one of the items changed. The external approach implemented in ATESS allows the analyst to concentrate on data relationships rather than on the mechanics of writing and validating source code, relieving the analyst of the overhead associated with developing software to read, process, and write data files from different sources.

## FUNCTION

The individual modules of ATESS can be used in a wide variety of combinations to produce different processing flows to meet the needs of different test programs. At one end of the spectrum, ATESS has a module designed primarily to read input data files written on "other" computers and produce output files in "standard" host computer files. At the other, ATESS includes a module designed to provide the analyst with a flexible data display capability (listings or graphics). Other modules perform such functions as data quality control checks; data file merging and subsetting under user control; event recognition based on time-correlated multiple data file contexts, data scaling, and validation; manual file creation and editing; and data archieval and retrieval. Each module is independent of the others, and there is no fixed sequence for processing data through the individual modules. Not all modules need to be used for a particular application, and each module may be used several times in a particular processing flow. ATESS lends itself to a building block approach, initially implementing a basic data flow to meet requirements at test start, then building to add multiple levels of

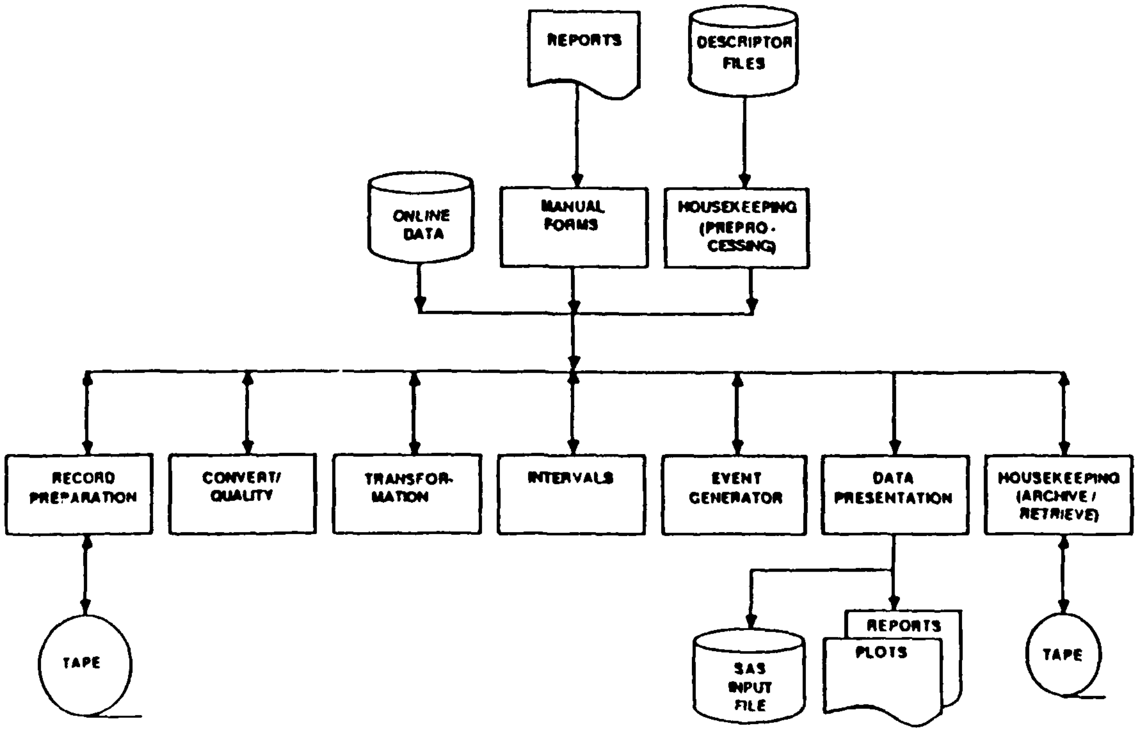
data analysis as the test progresses and data relationships are better understood. This flexibility is especially useful when analysis emphasis shifts during a test from planned to exploratory, as unexpected relations come to light, and new questions must be answered.

## STRUCTURE

ATESS is a modular set of software components that can be used in different combinations to perform data reduction and analysis functions to support T&E programs. ATESS consists of eight top-level functional components, encompassing approximately 800 routines or 40,000 lines of source code (Figure 1). Each functional component addresses one data processing function frequently performed in support of T&E. The eight components of ATESS are Housekeeping, Manual Forms, Record Preparation, Convert/Quality, Transformation, Intervals, Event Generator and Data Presentation, as shown in the overview. Each component provides the user/analyst with several capabilities, and data may be processed through as many iterations of the same or different components as required to attain the processing objective of the user. Several capabilities are available in more than one component, providing even greater flexibility. The order in which components are used for processing data depends on specific T&E program requirements and can be tailored to meet program-specific data type, format, and processing requirements using process control and data description files.

**Figure 1**

Analytical Test and Evaluation Software System Overview



## TOP-LEVEL COMPONENTS

Housekeeping - converts ACSII descriptor files into the required binary form for use in other system components. Using the VAX Backup/Restore command, this component also performs archieval/retrieval functions.

Manual Forms - is used to create new data files or to add, modify, or delete records in existing data files.

Record Preparation - extracts and translates raw data to user-specified parameter fields. It can be used to separate data into individual records, retrieve only required data from input disk or tape, and eliminate invalid data records.

Convert Quality - replaces or creates new data values by external specification. It can globally replace values, correct for calibrations, convert codes, or perform any user-defined function using external subroutines, relational operators, FORTRAN operators, and IF-THEN-ELSE statements. It also flags questionable data. The user defines the quality control check, the flag, and the replacement of the flag for each check. Questionable data are easily identified for further processing.

Transformation - restructures the data files and/or records without changing the data values. It combines multiple files into one file, or a file can be split into multiple files by splitting records in several ways: by outputting separate files for each unique value of a selected parameter, by selecting subsets of the original file, by sorting records, or by a combination of sorting and subsetting.

Intervals - eliminates redundant information from time-generated (vs. event-generated) data. The user can define a steady state, then specify one record for output containing start and stop times and any support data fields desired.

Event Generator - reads and interprets the contents of time-sequenced multiple input files, identifying user-specified data states. It performs the user-specified activities indicated in one or more work blocks associated with the event.

Data Presentation - produces listings, plots, and statistical summaries. A user-entered control file specifies the type, content, and format of the output. This user input has the flexibility to produce various types of listings, plots, and statistics using several different options.

## APPLICATIONS

The PLSS test team was the first user of ATESS components. The software was used to reduce, process, archive and retrieve data gathered during limited IOT&E of PLSS. Data flows through ATESS have been defined for F-4G WW PUP which is expected to begin testing in FY88. The remaining current users, Over The Horizon Backscatter Radar (OTH-B) East Coast Radar System (ECRS) and Tacit Rainbow (TR), have defined preliminary data flows and will be using ATESS during IOT&E in the near future. Additionally, Consolidated Space Operations Center (CSOC), High-Speed Anti-Radiation Missile (HARM) and MILSTAR test teams as well as (Tactical Warfare Center (TAWC) and Naval Weapons Center (NWC)), personnel have inquired into the use of ATESS. The number of potential users is unlimited because this same software could be applied to virtually any new test program, regardless of differences in data structures and methods of evaluation. Thus the use of ATESS can significantly reduce future software development efforts.


**Pat L. Brannen** is an analyst in the Modeling and Analysis Division at AFOTEC. She chairs the ATESS Configuration Control Board and is project officer for the ATESS Maintenance Subtask. Her former position was in the Nuclear Safety Department, Naval Weapons Evaluation Facility, KAFB.

**Ms. Kathleen Hibson** is manager of Configuration Control and Quality Assurance for The BDM Corp. She is currently the Program Manager for the ATESS Configuration Management Task, AFOTEC Subtask 447. Prior experience includes developing software systems for the Precision Location Strike System, the Space Transportation System, and the High Technology Test Bed with BDM. Ms. Hibson received her bachlelor's in Systems Planning and Management from Stevens Institute of Technology in 1982.

# Coping with Mission Support Factors in Test Planning and Evaluation

**By Samuel G. Charlton Ph.D.,**
**AFOTEC's Human Factors Division**

## INTRODUCTION

Mission support factors are those elements of a system typically considered during the course of a test, but which are not normally elevated to the COI objective level. These mission support (a.k.a. external) factors, contribute to mission effectiveness as necessary prerequisites without themselves defining mission success. Mission support factors include but are not limited to: in[illegible]ability, compatibility, training, safety, hum[illegible] security, and survivability.

The importance of these mission support factors has become increasingly apparent in our OT&E efforts. A significant percentage of service reports (SRs) are specific to the mission support factors mentioned above. In addition, these factors are receiving increased attention from DOT&E. Mission support factors are important contributors to system effectiveness and need to be evaluated as such. The object of this article is to briefly describe a promising approach to test planning and evaluation as we have applied it to the OT&E human factors of modern Air Force systems.

## FINDING THE TARGET

One of the principal problems inherent in evaluating mission support factors in modern Air Force systems is the distinctive and changing characteristics of these systems. Owing to their size, complexity, and cost, many of these systems, particularly in the strategic and communications arena, are one-of-a-kind, evolutionary systems. Unlike smaller products where a production decision can be based on extensive research and testing of mission support factors considerations, these large systems are unique and thus a decision to purchase one comes first, followed by expenditures of time, effort, and dollars to make it work. The development and test cycles are typically cut quite short to save money. As a result there is often too little time spent considering the demands that will be placed on operators and maintainers by the operational system.

Adding to this problem are the frequent modifications made to the system hardware, software, and procedures. The larger systems are dynamically reconfigured, the commands and procedures required to operate them are changed as their loads and missions are defined and redefined. The mission support factor considerations of the original system configuration may bear little or no resemblance to the mission support factor issues pertinent to the system after having undergone extensive engineering and procedural changes. The transitional nature of these systems also has the effect of making it difficult or impossible to provide adequate training for the operators and maintainers, there being no static system on which to base a training program.

## THE PROBLEM OF CRITERIA

A second, perhaps more fundamental problem inherent in evaluating and/or assessing mission support factors in OT&E is the lack of approved standards or criteria on which to base the evaluations. In some cases, technology advances have outpaced the ability of industry to define acceptable standards for software performance, operator performance, or man-machine interfaces. Similarly,, many system users are unable to specify criteria for these mission support factors until they have an opportunity to actually begin to operate the system. A related issue concerns the selection of which mission support factors to test and to what level of detail. To this end, it is readily apparent that one cannot consider these mission support factors in isolation. Mission support factors are important only in the context of their impact on system effectiveness. Further, the interactions between the software, hardware, human operators and their training programs must be examined simultaneously.

In the case of human factors, this issue arises *in the form of lack of human factors standards* or criteria on which to base the OT&E of complex military systems. While criteria do exist for the optimal physical layout and anthropometry of an operator's workstation,

these are not the limiting factors in operating large C3 systems. The pertinent human factors issues involve identification of the limits of human operators in assimilating and responding to the information presented to them. Aside from some very basic knowledge about human information processing and memory capabilities, few criteria exist for evaluating the complex interactions arising from multiple cognitive tasks. Nor are there criteria for determining the correct way to present the information required to support the diverse and changing repertoire of operator tasks.

## AN INTEGRATION APPROACH

How then are system evaluators to cope with the aforementioned problems? One promising approach we are currently employing in the OT&E of several major C3 systems is based on a mission support-system effectiveness integration approach to the identification of mission support factors "risks" in system design. The integration approach to human factors attempts to identify risk factors for the performance of C3 operators. In the absence of data specifying the causal relationships between system design, human information processing, and human error, this approach establishes a circumstantial case for the identification of situations and operator tasks that are "at risk" or closely associated with system failure. This approach is analogous to the medical epidemiological approach specifying risk factors associated with a disease in the absence of physiological data specifying the cause of the disease.

Briefly, the integration approach involves three procedural steps. The first step is the identification of which mission support factor measures to include in an OT&E plan. To maximize the efficiency and effectiveness of our test efforts, mission support factor OT&E measures should be selected on the basis of their relationship to system effectiveness rather than simply because they are measurable or have historical precedent. In essence, all our measures should have to pass the "so what?" test prior to their inclusion in an OT&E plan. That is, OT&E measures must either possess user-specified criteria, or be logically and quantifiably demonstrable as having an impact on the user's requirements for system performance.

For human factors OT&E, we identify candidate measures of human performance, human factors engineering and system performance. The human performance measures include operator response times, decision times, and response series times and error frequencies, thus representing several levels of task complexity. These human performance measures are drawn from the human-intensive components of system functioning such that they are representative of operator activities and relevant to system effectiveness.

The human factors engineering measures can include the results of human factors questionnaires administered to system operators or checklists completed by human factors engineers evaluating the system. These human factors engineering measures focus on issues of display formats, data entry procedures, system documentation, ambient noise and illumination and other issues relevant to operator performance. Measures of system performance include system throughput times, failure rates, or other measures related to the user-specified requirements for system effectiveness.

Once operational data of the above three types have been collected, the next step involves determining the degree of association between the mission support factors and system effectiveness via multivariate statistics. The key to this step is linking the criteria-less mission support factors measures to the objective criteria for system performance. For example, when system effectiveness measures are found to be marginal or failing, measures of human performance are ranked according to their degree of relationship with those system functions. Human tasks having a significant predictive relationship with system performance, as determined statistically, are identified as areas of human factors focus. The human factors engineering measures are then statistically related to those human performance tasks in order to determine the human factors issues having the greatest impact on the system operators' performance.

The results of interest are the human tasks identified as having significant relationships to system performance deficiencies and the human factors engineering issues possessing significant relationships with those human tasks. The criteria or measure of effectiveness for human factors engineering in these systems is thus operationally defined as a positive contribution

to overall system effectiveness. Candidate statistical methods range from the relatively simple multiple regression analysis to the more sophisticated factor analytic and categorical clustering procedures.

Finally, after the relative degree of association of the various human factors measures has been determined, one can attempt to derive explanatory or causal connections between operator performance and the human factors considerations of the system design and operations. These explanatory connections, can then be used to begin formulating more generic sets of human factors design criteria to be used in the design, test, and evaluation of future systems.

## SUMMARY

The integration approach to mission support factors test and evaluation is capable of addressing many of the problems identified earlier. The systems-level view of the approach ensures that our evaluations and assessments are first and foremost directed at addressing user requirements for system performance. All mission support factors to be assessed in OT&E are placed in the larger landscape of their impact on system effectiveness. The lack of human factors test criteria is also remedied by using overall system effectiveness as the measure of performance by which human factors issues are judged.

In other words, areas of human factors design are identified as being problems only inasmuch as they are statistically associated with poor system performance. Similarly this approach has the advantage of considering a variety of mission support factors simultaneously, thereby allowing the identification and evaluation of the interactions between different mission support aspects in terms of their relative contribution to system effectiveness.

The products of the integration methodology also provide a means for focusing remediation and corrective actions on those areas with the greatest impact. The specific mission support factor problems associated with a given system function are ranked in order of importance so as to provide a hierarchy of effective remedies. Further, the approach can provide periodic "snapshots" of dynamic mission support factors issues as they evolve in the frequently changing environment of a modern C3 system.

Finally, this approach can serve as a starting point for the development of a taxonomy of mission support factors criteria. Through the systematic use of this and similar evaluation techniques, an empirical database of mission support factors standards can be collected and generalized for application to the test and evaluation of new systems. Faced with a need to field effective operational systems, and in the absence of definitive human factors criteria, the integration approach described above has proved to be quite valuable in the dynamic environment of human factors test and evaluation.

**Dr. Samuel G. Charlton** is an Engineering Psychologist with the Aeromedical/Human Factors Division of AFOTEC where he functions as a human factors analyst for a variety of strategic and communications systems including: the Consolidated Space Operations Center, the Over the Horizon-Backscatter Radar system, the Global Positioning system, the Automated Remote Tracking System and related programs. He also serves as an Adjunct Assistant Professor of Psychology at the University of New Mexico. Prior to his assignment at AFOTEC, Dr. Charlton held positions with the BDM Corp. (Senior Staff Member), The University of Albuquerque (Assistant Professor), and the Veterans Administration Medical Center (Neuropsychology Research Associate). Dr. Charlton holds a bachelor's in Experimental Psychology from San Jose State University and a master's and Ph.D. in Experimental Psychology from the University of New Mexico. He is a three-time recipient of the National Institute of Mental Health (NIMH) National Research Service Award and was selected as UNM's Outstanding Doctoral Candidate of 1982 by Sigma Xi, the Scientific Research Society.

# Simple Linear Regression and Nonparametric Slope Estimation Techniques for Navigation Drift Rate Caculation using Non-Continuous Time, Space, Position, information (TSPI)

By Capt. Robert A. Eisenberger and
Capt. Christopher A. Strickland

## INTRODUCTION

The B-1B Follow-on Operational Test and Evaluation (FOT&E) test team evaluates the effectiveness of the B-1B weapon system and its supporting subsystems. Two such systems are the inertial navigation system (INS) and the dead reckoning (DR) system. Although these systems are common to many aircraft, the testing method used by the FOT&E team is unique due to the instrumentation systems restriction of non-continuous Time-Space-Position-Information (TSPI). The purpose of this paper is to describe two techniques to estimate the true drift rate of the navigation systems without a continuous TSPI source. We will discuss two methods of measuring the time history of navigation position error, two methods to estimate navigation drift rate from the position error history, and a Monte Carlo analysis which describes each techniques' robustness for different quality navigators, and both small and large sample sizes.

## POSITION ERROR MEASUREMENT

FOT&E uses two sources of measuring the position error of the navigation system. One method uses Time, Space, Position Information (TSPI) measurements of aircraft position from an instrumented range. These measurements are assumed to be truth and compared to aircraft recorded estimates of position. The difference between these measurements is the aircraft navigation system position error. The second method uses the aircraft's radar system. It measures a range and bearing from the aircraft to a Defense Mapping Agency (DMA) surveyed ground point. This range and bearing estimate is added to the known position of the surveyed ground point to calculate aircraft position. It is then stored in the aircraft's computer complex and compared to the navigation system position. The difference between the two values is automatically calculated and stored in temporary computer memory called the "radar buffer."

## TIME, SPACE, POSITION, INFORMATION (TSPI)

In discussing method one, a description of the aircraft instrumentation system must first be presented. The Production Data Acquisition system (PDAS) is a palletized instrumentation system which is installed in the B-1B Central Equipment Bay. The pallet contains a processor chassis, Power Distribution Unit, and a disk drive assembly. The Disk Assembly is removable and can hold up to $\pm$ 160 megabytes of data. The PDAS has an internal clock which is initiated with a time synchronization (WWV) signal and maintains time throughout flight within 60 milliseconds per eight hours. The processor receives inputs from up to six MIL standard 1553B avionics bus sources, reformats and edits the data, and passes it on to the removable disk for storage. Following a test mission, the disk is removed from the aircraft and transported to a ground station where the data is extracted from the disk for further processing. Using the PDAS, INS and DR latitude and longitude are recorded continuously throughout flight.

The PDAS is a passive listening device and cannot act as a TSPI source. It does not have an independent reference such as the Navigation Instrumentation Accuracy System (NAIS) or the Completely Integrated Reference Instrumentation System (CIRIS). FOT&E plans each mission to pass through the radar coverage of multiple instrumentation ranges along the route of flight. Coverage over each range lasts from 2 to 30 minutes depending on mission profile, range availability and range capabilities. Due to the limited number of ranges available on any given day, we receive one to four short periods of continuous TSPI and long periods of no TSPI. A position error time history of these portions of continuous TSPI coverage is calculated post-

flight. The calculation uses an oblate (assumes the earth is elliptical) earth model to convert differences in latitude and longitude to differences in feet (3).

The errors associated with using this technique of position error time history measurement are as follows:

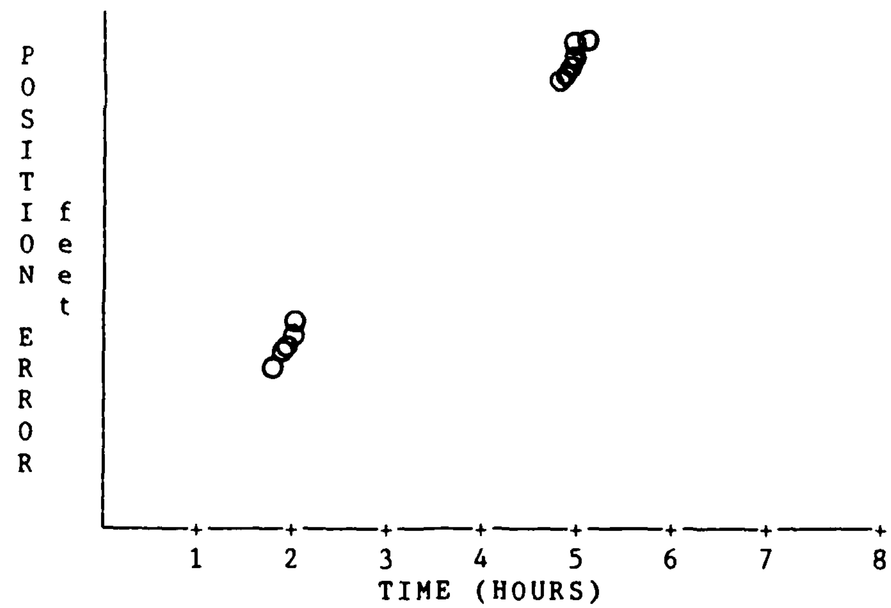
Timing difference between the PDAS clock reference and range time reference. Both sources synchronize to WWV. The range uses an interrange instrumentation group (IRIG) time generator to maintain the reference, while the PDAS uses an internal crystal. Since the PDAS can drift up to 60 milliseconds after eight hours, a position error is created due to the aircraft position and site alignment inaccuracies. Though actual capability varies from site to site, an estimate of 100 feet of error bounds all systems.

Oblate ellipsoid model assumes the earth is an ellipse with a specified eccentricity. Though this is not totally accurate, for short distances (< 20 miles) this error is small. FOT&E testing shows it to be approximately 10 feet.

Therefore, summing the values from a,b, and c above, the error in measuring position error by comparing aircraft recorded position to range measured position is on the order of plus or minus $\pm 150$ feet.

An example of the data accumulated using this technique is plotted in Figure 1.

Figure 1




airspeed during this 60 millisecond lapse. With an average cruise speed of 750 ft/second, typical errors due to time are approximately 45 feet (750 ft/sec * .06 seconds).

Range measurement error is the ability of the instrumentation range to measure aircraft position. This error includes errors due to tracking the aircraft at other than the INS

## RADAR BUFFER READINGS

As stated earlier, the second method of accumulating a time history of navigation position error uses the aircraft radar system. In Figure 2, the navigator commanded a radar map of an area that contains a surveyed point. For this example, the point is the center of a bridge

spanning a river between two land masses. Each point on the radar map has a latitude and longitude associated with it as measured by the INS. By overlaying a set of moveable computer generated crosshairs on the map, the navigator has the ability to determine his position error. The procedure follows:

**a.** The map is generated as in Figure 2 when the navigator commands the INS to place the crosshairs on what it thinks should be the aimpoint.

**b.** In the example, the INS placed the crosshairs northwest of what the radar system shows to be the aimpoint (center of the bridge).

**c.** Since the INS commanded (uncorrected) placement of the crosshairs is in error, a position error exists in the INS.

**d.** The navigator moves the crosshairs on the aimpoint through a hackhandle control stick with the result being Figure 3.

**e.** On the lower left corner of the map, the radar buffers are displayed. The top number represents the error in the INS in distance, while the bottom number represents the bearing. These buffers are accurate for the time of map generation, displayed on the lower right corner of each map. These numbers are only accurate when the crosshairs are on the aimpoint. (Figures 2 & 3)

This information is hand recorded on a test card every 15 minutes by the navigator. After assuming that the Schuler frequency is the highest frequency in the signal, a sampling rate can be determined. Fifteen minutes of sampling is required to satisfy Shannon's sampling theorem which states "A band limited signal can be uniquely represented by a set of samples taken at time intervals spaced less than 1/2 W seconds apart, where W is the highest frequency in the signal [5]." Since the frequency of the Schuler period is 84.4 minutes, Shannon requires samples to be taken at least every 42 minutes.

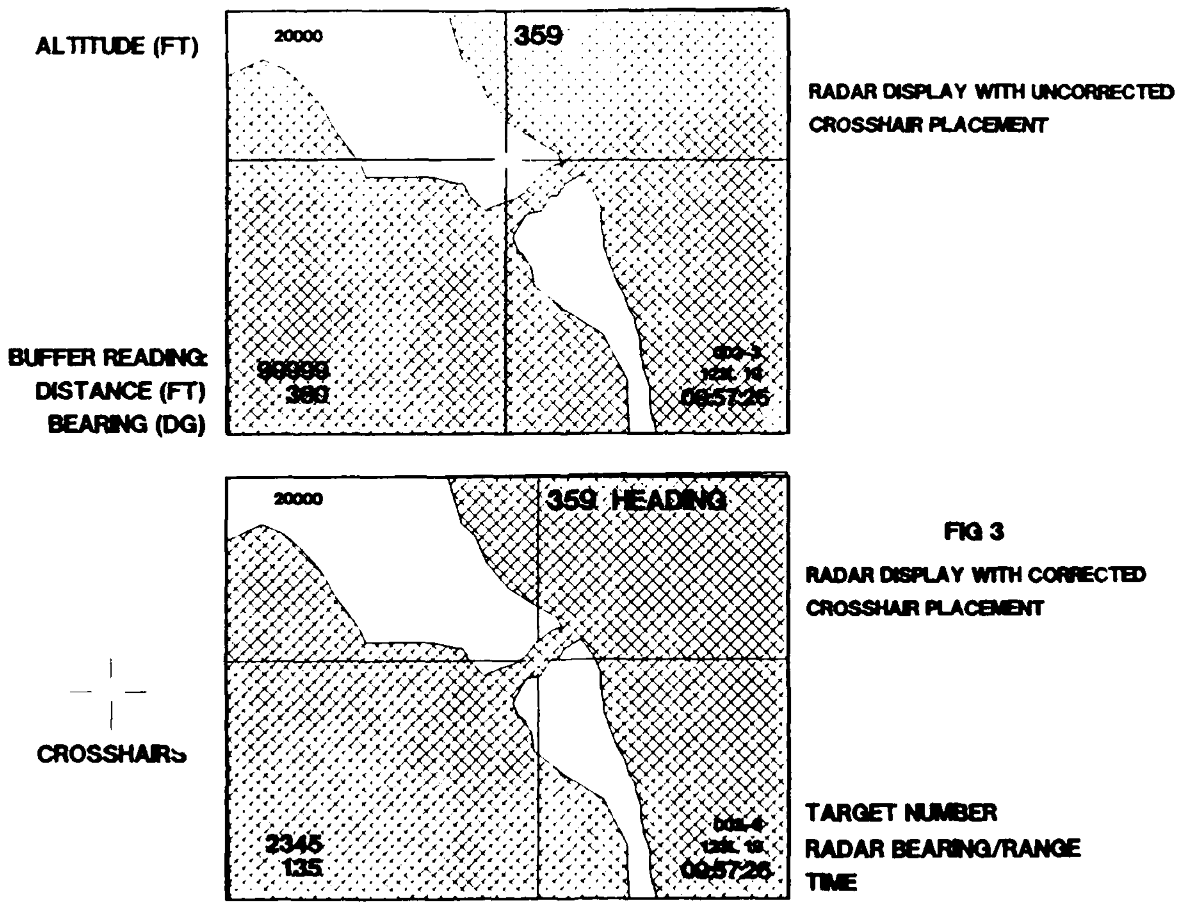
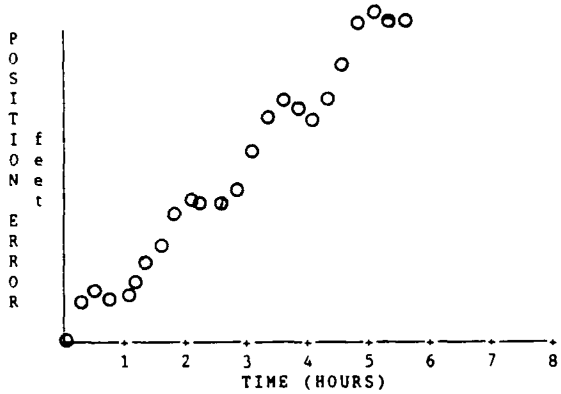



FIG 3

Using this technique, a position error time history plot can be created that has one sample every 15 minutes throughout the mission. An example is plotted in Figure 4.

Figure 4




The errors associated with using this technique of position error time history measurement are as follows:

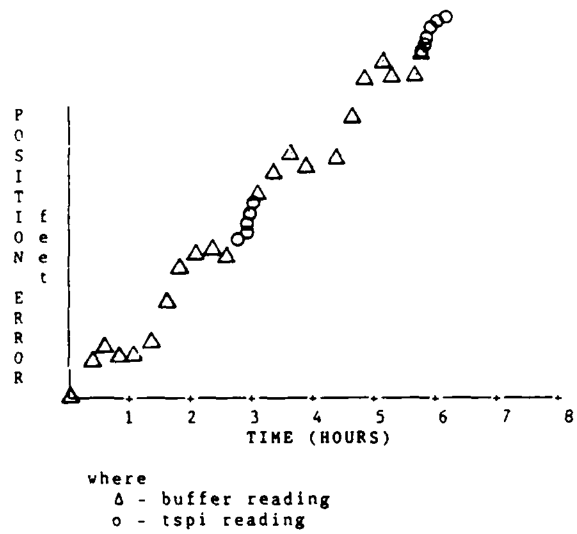
<u>Avionics operator crosshair placement error</u> - This is the ability of the navigator to identify an aimpoint on the radar display unit and position the crosshairs exactly on the top of it. It also includes the error due to picture distortion, beam smearing and antenna misalignment. This error has been measured through flight test and is shown to be consistently less than 200 feet [6].

<u>Target survey measurement error</u> - The Defense Mapping Agency publishes measurement errors associated with each aimpoint. Aimpoints with 50 feet error or less are used for our analysis.

## COMBINING TECHNIQUES

Whenever possible, a combination of the two position error measurement techniques are used to accumulate a position error time history (see Figure 5).

Figure 5




## DRIFT RATE ESTIMATION

## SIMPLE LINEAR REGRESSION (SLR)

The most common method of estimating a drift rate is to use *the slope $\hat{\beta}_1$ from the SLR model:*

(7) $$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 t + \varepsilon,$$

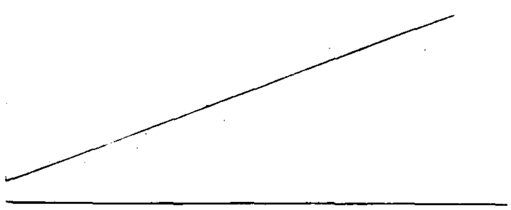
where $\varepsilon \sim \Phi(0, \sigma^2)$.

| | | |
|---|---|---|
| In (7) $\hat{y}$ | - | the fitted data point. |
| $\hat{\beta}_0$ | - | the y intercept. |
| $\hat{\beta}_1$ | - | the slope. |
| $\varepsilon$ | - | the difference between the fitted and the actual value. |
| $\Phi(0, \sigma^2)$ | - | normal distribution with a mean of zero and some unknown variance ($\sigma^2$). |

SLR is a method of fitting a linear model through a set of points. The model is fit by the method of least squares. In least squares

regression, the model is the line with the smallest squared deviations between points on the regression line and the actual data points (see Figure 6).

Figure 6
Simple Linear Regression Least Square Fit



## THE THEIL STATISTIC

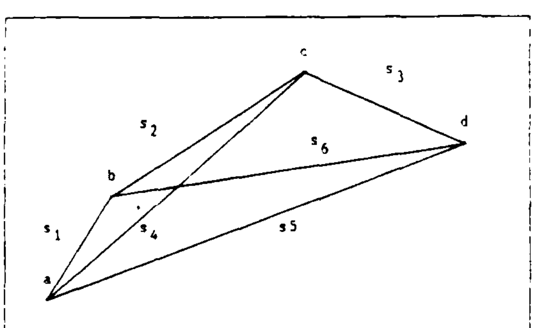
The next method is a better estimator of drift rates when the sample sizes are small and the data is erratic. It is called "An Estimator Associated With the Theil Statistic" by Theil. The model built is similar to the model used in equation (7). However, normality of the errors is not required. The only assumptions are as follows:

**a.** The errors $\varepsilon$ are mutually independent.

**b.** Each $\varepsilon$ comes from the same continuous population.

This method has practical test applications when estimating drift rates since it uses all possible slopes (see Figure 7).

Figure 7
Slope Estimated From The Theil Statistic




This method tak[illegible] of all of the "what if" situations. For example what if I started at point b and stopped at point d, what would my drift rate be? The Theil Statistic for Figure 7 is the median of slopes $\{s_1, s_2, s_3, s_3, s_4, s_5\}$. It should be noted that the estimator of slope from the Theil statistic is less sensitive to gross errors than is the classical least-squares estimator from the SLR method. This is because the Theil Statistic takes the median of the slopes, whereas the classical slope is computed by a weighted average of the slopes [7].

## MONTE CARLO ANALYSIS

To determine the robustness of each drift rate estimation technique, a Monte Carlo analysis using 4800 samples was performed. This analysis varied as to

**a.** Schuler amplitude

**b.** Drift rate

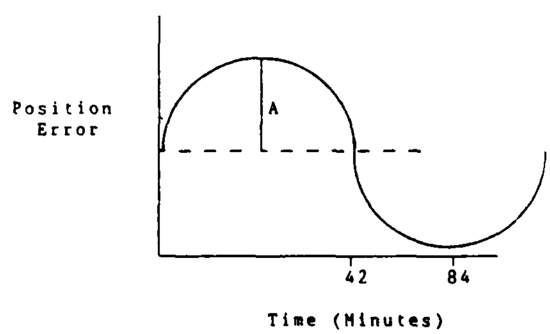
**c.** Position error sample size

Each of these parameters were first varied separately, i.e. Schuler amplitude was varied while drift rate and sample size remained constant. Additionally, two variables were varied while holding the third constant. This approach allowed us to determine the ability of each estimation technique for different types of navigators and different data collection systems.

## SCHULER AMPLITUDE

The first variable investigated was Schuler amplitude (4). A model depicting this 84 minute error is shown in Figure 8.

Figure 8
Schuler Tuned Inertial System Error



The amplitude "A" above varies with different quality navigators or aiding techniques. The lower the platform misalignment error (better the INS), the lower the Schuler amplitude. For aided systems (i.e. doppler aided, stellar) misalignment error is periodically reduced, thus decreasing the Schuler amplitude.

## DRIFT RATE

The model incorporated different drift rates with these values representing systems with strategic accuracy, such as an ICBM, through coarse navigation schemes used in civil aviation.

## POSITION ERROR SAMPLE SIZE

This parameter was varied to estimate realistic sample size opportunities. This sample size was selected randomly from the 360 position points generated from the six hour navigation leg. Though, one sample every 15 minutes is the minimum recommended sample from the total population, we will show smaller sample sizes can be used to estimate an accurate drift rate.

## RANDOM ERROR MODEL

The simulation involved 4800 six hour navigation legs. Each time a nav leg was generated, it computed one sample every minute for a total of 360 samples per navigation leg. The nav legs were based on a Shuler with amplitudes of 250, 500, 750, 1000, 1500 and 2000 feet and drift rates of .1, .3, .5, .75, 1, 1.5, 2 and 3 nm/hr. From each nav leg random samples of 5, 10, 15, 25 and 30 were drawn from the 360 points in the navigation leg. Twenty nav legs were generated at each amplitude, drift rate and random sample size grouping. The results of each of the groupings can be seen in Tables 1 - 3.

The formula for determining the model is as follows:

(8) $f(\tau, \delta : t) = \tau t + \delta \sin(.0744t) + \varepsilon(t)$

where

(9) $\varepsilon(t)$ is defined as follows:

if $U(0,1) < .95$ then

$\varepsilon(t) = U(-300,300)$

else

if $z = 150 \cdot \Phi(0,1) \geq 0$ then

$\varepsilon(t) = 300 + z$

else

$\varepsilon(t) = -300 + z$

$\Phi$ is a normal distribution

$U$ is a uniform distribution

$\tau$ is the shuler amplitude in feet.

$\delta$ is drift rate in feet per minute.

$t$ is time in minutes.

$f$ is the drift in feet.

$\varepsilon(t)$ is the model error

## RESULTS

Overall Comparison: Theil Statistic vs SLR. An analysis of variance (ANOVA) was applied to all of the test conditions against the following hypothesis:

$H_0$ - There is no difference between the slope in the Theil Statistic and the slope from a simple linear regression (SLR) for estimating drift rates.

$H_a$ - The slope from the Theil Statistic is a better estimator of drift rates.

Based on the results of the ANOVA, reject $H_0$ in favor of $H_a$. Over all three factors investigated an estimated drift rate based on the Theil slope gives you an error of .0032 nm/hr; while the SLR slope gives you an error of .0046 nm/hr.

By Sample Size: Theil Statistic vs SLR. From Table 1 it can be seen that in most cases the Theil Statistic was a better estimator of drift rate than the SLR technique, though the two estimators tend to converge as the sample sizes increase.

Table 1
Error in drift rate estimation by Sample Sizes
based on 960 nav legs per sample size

| Sample Sizes | Theil Error Mean | SLR Error Mean | Theil-SLR Mean | PValue |
|---|---|---|---|---|
| 5 | 0258 | .0323 | -.0063 | .0001 |
| 10 | -.0009 | -.0016 | .0007 | .0001 |
| 15 | -.0126 | -.0129 | .0003 | .0242 |
| 25 | .0047 | .0057 | -.0010 | .0001 |
| 30 | -.0013 | -.0006 | -.0007 | .0001 |

By Drift Rate: Theil Statistic vs SLR.

Table 2 shows the Theil Statistic to be a better estimator of drift rate than the SLR technique. The error between the two techniques is constant over all drift rates indicating the accuracy of the slope techniques is unaffected by this variable.

---

†Pvalue is the smallest probability that we can reject the hypothesis that the Theil slope equals the regression slope. (Typically we would reject this hypothesis for any Pvalue < .05)

Table 2

Error in drift rate estimation by Drift Rate based on 600 nav legs per drift rate

| Drift Rate | Theil Error Mean | SLR Error Mean | Theil-SLR Mean | Pvalue |
|---|---|---|---|---|
| 100 | .0034 | .0048 | -.0014 | .0001 |
| 300 | .0035 | .0050 | -.0015 | .0001 |
| 500 | .0027 | .0048 | -.0016 | .0001 |
| .750 | .0034 | .0049 | -.0015 | .0001 |
| 1.000 | .0029 | .0043 | -.0014 | .0001 |
| 1.500 | .0029 | .0043 | -.0014 | .0001 |
| 2.000 | .0032 | .0044 | -.0012 | .0001 |
| 3.000 | .0032 | .0047 | -.0015 | .0001 |

By Shuler Amplitude: Theil Statistic vs SLR
As in the other two tables, the Theil Statistic is the better estimator of drift rate. There are noticeable trends in the difference between the two methods as the Shuler Amplitude increases. With an increase in amplitude, both estimation techniques increase their error; however, the SLR technique is increasing at a higher rate.

Table 3

Error in drift rate estimation by Amplitude based on 800 nav legs per Schuler Amplitude

| Amplitude | Theil Error Mean | SLR Error Mean | Theil-SLR Mean | Pvalue |
|---|---|---|---|---|
| 250 | .0005 | .0010 | -.0005 | [illegible] |
| 500 | .0014 | .0024 | -.0010 | [illegible] |
| 750 | .0024 | .0038 | -.0014 | [illegible] |
| 1000 | .0028 | .0043 | -.0015 | [illegible] |
| 1500 | .0052 | .0070 | -.0018 | [illegible] |
| 2000 | .0065 | .0092 | -.0027 | .0001 |

## CONCLUSIONS

The B-1B Follow On Test and Evaluation Test Team must perform navigation effectiveness testing without a continuous TSPI source. Two methods were presented to measure the navigation error time history of the INS. One method uses an on-board recorder to record INS position and compare it to non-continuous TSPI sources to calculate the error. The second method uses the aircraft radar system as a direct measurement of the error. From these time histories, drift rate is calculated by using an estimator of slope. Two techniques were presented to derive this estimator. The first method, simple linear regression, proved to be accurate within .0046 nm/hr average error. As sample rate increases, performance of this estimator gets better, however the more the Schuler cycle is allowed to remain undamped, the poorer the performance. The second method used the Theil statistic to estimate drift rate. It proved to be a more robust estimator under virtually all test conditions. Its average error is .0032 nm/hr.

## REFERENCES


1. Offensive Avionics Production Data Acquisition System (PDAS) Users Manual, Rockwell International, Document # NA-86-1388, 17 October 1986.

2. Production Data Acquisition System (PDAS) users Manual, Boeing FSCM No. 82918, Document # D400-45001-7, 23 April 1986.

3. U.S. Coast and Geodetic Survey Special Publication Number 241.

4. American Practical Navigator, Pub. No. 9, Bowditch, 1977.

5. Continuous and Discrete Signal and System Analysis, McGillem and Cooper, Holt, Rinehart and Winston, Inc., 1974.

6. Monthly Operational Suitability and Effectiveness Summary, B-1B Follow-on Operational Test and Evaluation Test Team, 30 November 1987.

7. Myles Hollander and Douglas Wolfe, Nonparametric Statistical Methods, John Wiley and Sons, 1973.

8. Sanford Weisberg, Applied Linear Regression, John Wiley and Sons, 1980.

**Capt. Robert A. Eisenberger** is an astronautical engineer graduate from the United States Air Force Academy. He holds a bachelor's and master's in Engineering and an M.B.A. in Aviation. He was previously assigned to the Ballistic Missile Office as a project engineer for advanced re-entry vehicles.

**Capt. Christopher A. Strickland** is a master statistician graduate from the Florida State University. He holds a bachelor's with majors in Mathematics and Statistics and a master's in Applied Statistics. He was previously assigned to the Office of the Assistant Secretary of Defense (Comptroller) as a software engineer for budget analysis.

Both officers work for the Data Management Deputate of the B-1B Follow-on Operational Test and Evaluation Test Team at Dyess AFB, Texas. The Test Team's objective is to determine the operational effectiveness and suitability of the B-1B.

# Tactical Operational Test and Evaluation

By Mr. Donald L. Giadrosich,
USAFTAWC Chief Scientist

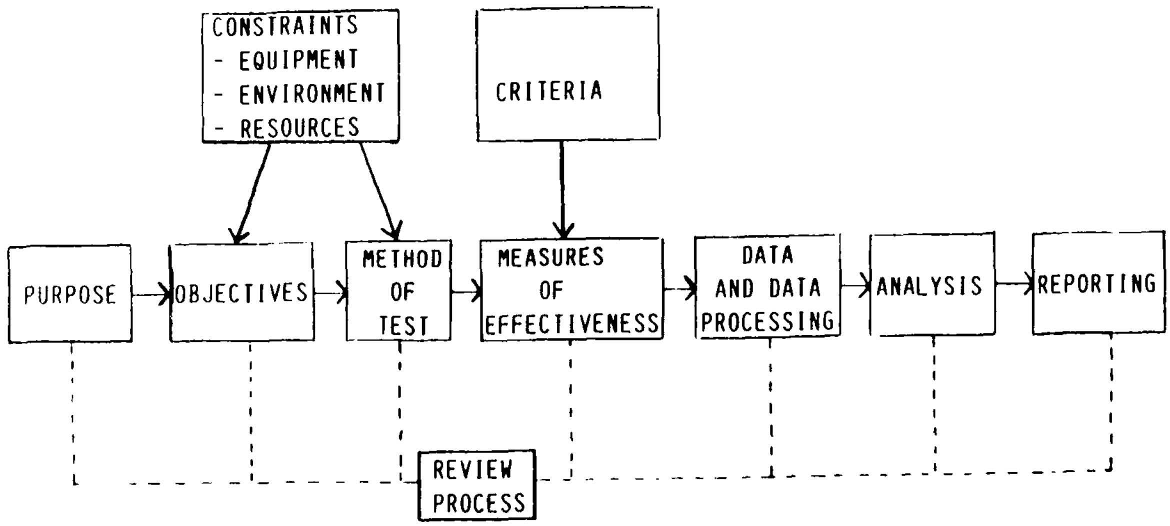
## INTRODUCTION

Operational testing can be defined in many ways, but central to all of the definitions is the acquisition of information to support acceptance or rejection of an idea, concept, or weapons system. Operational testing establishes or reinforces the confidence of decision makers that a system is indeed ready for production or application in combat. For those systems that are judged to be ready, it refines their performance by identifying what can be done to make them even better, and for those that aren't ready, it should explain why and whether they can eventually be made ready. For most complex tactical systems, acceptance or rejection is rarely a clear cut issue totally supportable by quantitative data, and the informed judgment of the operator must play a vital role in the decision process. This article focuses upon some to the difficult tradeoffs made in operational testing and the important roles played by the operational users, testers, and developers.

## SYSTEMATIC APPROACH

For over 20 years tactical operational testers have used a systematic approach in the design and execution of operational tests and tactics development projects. Figure 1 depicts a simplified version of this approach. We start with a purpose that is amplified by the objectives. The method of test is selected based on how the idea, concept, or system relates to the appropriate tactical air missions and the expected threat environments. For each objective, one or more measures of effectiveness are chosen to assess performance with regard to that objective. Both qualitative and quantitative measures are important and are used with emphasis upon operational realism, relevance, sufficiency, accuracy, and achievability.

FIGURE 1
SYSTEMATIC APPROACH TO TACTICAL
TESTING/TACTICS DEVELOPMENT



The measures of effectiveness lead to the data that must be collected during the test and to the instrumentation and methods that need to be used in collecting, processing, and summarizing the data for analysis. When decisions are made in a test report, the test report should define the criteria that explains the basis of the decisions. Criteria, in turn, should be based directly upon the operational need which considers both friendly and enemy capabilities. Since most systems require a long time in development and capabilities change with time, the criteria must reflect the latest information to ensure that the system produced can meet the threat. The most appropriate time and place for final commitment to criteria is after the decision maker has had the benefit of the knowledge and experience gained during operational testing.

## TEST DESIGN ISSUES

In today's complex testing world, the operational tester always faces the problem of trying to decide what to physically test and what to simulate. On one hand, there is the urge to be operationally realistic by dropping bombs, shooting bullets, and so forth. On the other hand, computer simulations can examine many situations that can't be physically tested, and often simulation can be done at significantly less cost. In some cases, a combination of physical testing and computer simulation is required to examine the full spectrum of issues, threats, etc.

In my view, there is no substitute for physical operational testing if it is within the realm of being possible and practical. In those cases where it is not possible for physical operational tests to satisfy the issues, computer simulations and other analytical practices may be in order. However, the bottom line is that if you can fly it and test it, that's the most desirable and credible process. The operator needs the assurance provided by physical testing that the actual weapon system (i.e., not the computer simulation) will work given that he has to employ it in combat.

In any test design, there is always the question on how much data are enough. One answer to this question is probabilistic, in that it has to do with how much risk the decision maker is willing to accept regarding the "proof" or "disproof" of the idea, concept, or system. A confidence level of at least 80% should be demanded by any prudent operator where quantitative measures are applicable and the risk can be quantified. Regardless of whether quantitative or qualitative measures are being used, the sample size must be large enough to convince the operational expert that the results are valid. If this is not possible, then the tester needs to find more credible ways and measures to make the evaluation.

Operational testers should be highly sensitive to the need for the right balance of instrumentation versus no instrumentation. If the operational tester can understand what's going on by relying upon operator observations augmented by a minimal amount of instrumentation, that's the way it should be done. Any time instrumentation is introduced into the measurement process, there is also an increase in the cost and time involved in the test, a corresponding risk that the measurement system may adversely affect the environment, and a tendency of the tester to rely upon the numbers generated rather than tactical operational judgment. Too little instrumentation can result in not understanding what's going on in the test; whereas too much instrumentation may bias the results and/or make it impossible to complete the test at a reasonable cost and within a reasonable time frame.

## OPERATIONAL TEST INDEPENDENCE

A lot has been said in recent years about the need for more independence in operational testing. This immediately should lead one to the question: independence from what? On a new system, the developer beyond question has a vested interest in "making the system work" and he also has the most knowledge about "how the system works." As taxpayers and interested citizens, most of us are reluctant to see a military system scrapped after several years and millions of dollars have been expended in its development. On the other hand, whether developer, tester, or user, we must demand and get a system that satisfies the current needs of the user. Anything less is shirking our responsibility and is not acceptable.

There also has been a rather loose use of the term "operational user." The operational users clearly are the operational commands which will eventually receive the systems and take them

into combat if the need arises. There may be a reason for some independence by the tester from the developer to ensure objectivity in the evaluation; however, arguments regarding tester independence from the operational user is like trying to select the automobile you like for your neighbor.

## SUMMARY

Acquisition and initial testing of weapons systems today is more of a team process than it ever has been in the history of the Air Force. The savvy of the developer, the independence of the test command, and the operational expertise of the user are all essential. The informed judgment of the operator based on physical test data and experience has always been and still is the most powerful ingredient of operational testing. Computer simulations and other forms of analysis are used where appropriate to augment and supplement the results of physical tests.

The operational commands are the primary organizations responsible for stating requirements that lead to ideas, concepts, and systems. They play an important role in conducting operational tests and tactics development and in supporting the operational test and evaluation of major new systems. Operational commands will take the systems into combat if we go to war, and they will win or lose depending upon their effectiveness. Consequently, they have a high degree of interest in all systems and must always play a highly active role in the acceptance or rejection process.

**Mr. Donald L. Giadrosich** is the Chief Scientist/Chief of Operations Analysis for the USAF Tactical Air Warfare Center, Eglin AFB, Fla. He has served for over 31 years as a Senior Electronics Engineer in industry, Supervisory Research Analyst with the US Navy, Scientific/Technical Advisor with the Joint Chiefs of Staff, Chief Scientist with the USAF Tactical Fighter Weapons Center, and Director of Operations Analysis, Headquarters US Air Forces Europe. He holds a bachelor's degree in Electrical Engineering, a master's in Economics/Systems analysis, and numerous awards which include the Florida AFA General Lewis H. Brereton award and the Presidential rank award of Meritorious Executive.

# AFEWES and REDCAP: Indoor EC Ranges

By Lt. Col. Al Bryson, AFOTEC/OL-AI

## INTRODUCTION

Capt. Golden Arm is one of the Air Force's top F-16 drivers stationed at Briarpatch AFB. Today, however, he is strapped in the cockpit of an F-16 flying over enemy territory. His Airborne Self Protection Jammer (ASPJ) is interfaced with his Radar Homing and Warning Receiver (RHWR), and both are in the green. Suddenly, the Surface-to-Air Missile (SAM) radar that has been electronically painting his aircraft locks on, generating that familiar, yet hair-raising tone announcing danger. He checks the alphanumeric on the RHWR receiver showing the threat at one o'clock. The ASPJ electronically counters the threat, causing the SAM operator on the ground to activate various ECCM modes to reacquire the F-16 target. Suddenly a missile launch indication is received the F-16 cockpit. Golden frantically searches at one o'clock for a visual on the missile as he begins maneuvers. The SAM operator skillfully maintains track on the jinking target to keep the missile on its deadly intercept course. Suddenly, Golden has a tally on the streaking missile and in desperation attempts a last maneuver to avoid impact...

No, this is not a paragraph out of Red Storm Rising. This is a possible scenario from the Air Force Electronic Warfare Evaluation Simulator (AFEWES), located at Air Force Plant #4 (General Dynamics), Ft. Worth, Texas. I could have just as easily described an ingress of 100 Blue aircraft into Eastern Europe, and the reaction that the Integrated Air Defense System (IADS) would take to counter the invading force. That would represent a typical scenario from the Red Capabilities (REDCAP) test facility located at the CALSPAN Corp., Buffalo, N.Y.

## GENERAL HYBRID STRENGTHS

AFEWES and REDCAP are owned and managed by ASD RWW. ASD and AFOTEC maintain on-site operating locations to provide assistance to both developmental and operational testers. The operation and maintenance of the simulators are performed by contractor personnel. Both facilities use hybrid, electronic combat simulations. What's a hybrid? In simple terms it means that part of the simulated scenario is accomplished with real equipment, in real time, with a man-in-the-loop. The remainder of the simulation (RCS, RF in free space, etc.) is modeled in real time. The resulting hybrid simulation has both advantages and disadvantages when compared to range testing, such as the test range at Eglin AFB, FL. Some of its advantages are:

Security - Since emissions are contained with a secured and controlled environment, there is no opportunity for compromise of sensitive information. Therefore, SAR (and SCI at the AFEWES) data can be, and routinely is, used.

Repeatability - Scripted encounters can be repeated numerous times under the same set of test conditions. This allows the tester to systematically alter test variables under a controlled scheme to do sensitivity testing or technique optimization.

Scoring/Instrumentation - Each facility is easy to instrument due to the very nature of hybrid testing. A variety of MOEs can be selected to meet the test criteria. Real-time measurements, at many points of the signal path or C3 net, can be made and recorded on strip charts or stored in computer memory for display in almost any format the tester desires. Each scenario run produces a vast amount of data, and many runs can be made each hour of testing.

Pre-Fabrication Testing - Ideas for proposed systems can be evaluated against specific threat systems or scenarios to assess potential prior to large hardware development investments. Also, nonflightworthy equipment, such as brass boards, can be evaluated.

Signal Density - Both facilities can generate large numbers of RF signals providing a realistic background for equipment under test. Most EC systems have to process and make decisions on hundreds of thousands of pulses each second.

Hybrid simulators are the only places that realistic pulse density tests can be conducted.

Of course there are disadvantages of hybrid testing. The main ones are that the equipment under test is not in adverse flight conditions and the threat simulation is not transmitting into free space. These are areas where open air ranges are very capable.

## COMPLEMENTING OPEN AIR RANGES

Now the natural thing to do is to hold the hybrid lab and the test range up side-by-side and do a comparison to determine which is the single best place to test. Let me suggest that we not compare their individual capabilities, but that we look at how they compliment each other. Although the relative ratings on the chart below can be debated, it shows that where one type of test capability is weak, the other has strength. Therefore, the logical way to test EC equipment is not to decide whether to take it to a hybrid or a range, but to determine how to take advantage of the strengths of each by planning an integrated test.

| TEST FACTORS | HYBRIDS | RANGES |
|---|---|---|
| Availability | Very Good | Poor |
| Instrumentation | Very Good | Fair |
| Tactics Development | Very Good | Fair |
| Relative Cost | Good (medium) | Fair (high) |
| Signal Density | Good | Fair |
| IADS | Good | Fair |
| Environment Realism | Fair | Good |
| Operator Realism | Fair (one sided) | Very Good (two sided) |
| Hardware Interaction | Fair (subsystem) | Very Good (system) |
| Tactics Evaluation | Fair | Very Good |

## AFEWES VS. REDCAP

I have referred to both the AFEWES and REDCAP as hybrids, however, they are not redundant capability. As in the examples at the start of this article, AFEWES' primary mission is to test real ECM equipment on board designated blue aircraft against a point defense such as a SAM, AAA, or Airborne Interceptor. REDCAP's mission is to look at a broader picture. How does an aircraft, or multiple aircraft, penetrate an IADS? Where are the reporting delays, and what is the value of jamming EW GCI assets or cutting a communication link? Typically, results from AFEWES (or range) testing can feed directly into the REDCAP test.

## AFEWES CAPABILITIES

AFEWES has high fidelity, closed loop simulations of many of the wartime threats to aircraft of both US and friendly foreign forces. That is why many US Army, Navy, and foreign testers join the Air Force in doing RD, DT, and OT&E at the AFEWES. Each of the closed-loop simulators typically has integrated transmitter, receiver, signal processor, display and weapon elements in hardware. These simulators are designed to do effectiveness evaluations and have complete end-game scoring capability. That is, they can give missile artillery missed distance in real-time and Pk via off line modeling.

The AFEWES simulators have a great deal of flexibility, allowing a tester to ask "what if" questions. Such questions as "what if the Red system was mounted on a 30 meter tower rather than at ground level?" "What if the channels on the threat receiver had a three db imbalance due to needed maintenance?" "What if that was a six db or two db imbalance?" "What if I reduced the RCS on selected Blue aircraft?" This type of questioning can save time and money for the tester who knows the right question to ask and how to interpret the answer.

Most of these simulations have been through an FTD simulator validation (SIMVAL) and, undergo periodic upgrade, to align them with the latest intelligence estimates. As with all simulators, there are differences between the current simulators and the latest intelligence. A simulator validation does not validate that the simulator is exactly the same as the real threat. A SIMVAL report does, however, tell a tester the differences and explain the potential impact on testing.

AFEWES offers several other significant test capabilities to support the above simulators:

Multiple Emitter Generator (MEG) - Capable of generating up to 205 emitters at 175 sites. It can generate any signal in the .5 to 18 GHz range. The MEG is the heart of the AFEWES capability to test the modern power managed EC systems.

Jammer Technique Simulator - JETS can simulate most types of ECM in the 2-18 GHz range.

IR Optical Tracking - RF tests comprise the majority of the AFEWES testing, however, there are capabilities in other frequencies.

C2/C3 Simulation - Simulates brigade, battalion, and regimental headquarter control over some ASM battery simulations. Also, can evaluate RF jamming effects on voice and data links.

Flight Simulator Lab - This ultra-high visual resolution domed flight simulator is the property of General Dynamics. However, it is integrated with the AFEWES to do interactive tactics development and Blue maneuver evaluation.

## REDCAP CAPABILITIES

REDCAP, like the AFEWES, employs man-in-the-loop, real-time simulation. A simplistic comparison of AFEWES and REDCAP would show that AFEWES C3 capability supports its threat radar simulators, and REDCAP's EW/GCI radar simulators support its C3. The focus of REDCAP testing is to evaluate the penetration of area defenses. This can involve employment of tactics, ECM, force mixes, and/or C3CM to determine the best mix of EC elements within a total force structure. REDCAP is generally very useful whenever a penetration concept affects more than one element of an air defense system, where there are a number of independently acting penetrators, or where there are a multitude of simultaneous, coordinated threat signals present.

The 1986 addition of a SUAWACS simulation extends the EW/GCI capability of REDCAP to the air. This added dimension allows the REDCAP testers to explore questions that can not be addressed at any other test facility or range in the U.S.

The effects and actions in a large scale netted, multi-radar, multi-penetrator scenario are extremely complex. Therefore, a complete time history of significant parameters and events is recorded for later study and test analysis. The table below is a partial list of the typical measures of effectiveness that are used in REDCAP reports:

EW Radar and Reporting MOEs:

- -Probability of initial target detection,
- -Tracking accuracy,
- -Reporting rate, frequency, delays,
- -False target reports, number, duration.

Processing Center MOEs:

- -Track accuracy versus number of tracks,
- -Ability to correlate redundant tracks,
- -Delay from first report to established track,
- -Tracking accuracy through triangulation,
- -Ability of the net to correlate sparse data to direct weapons.

Weapons Direction MOEs:

- -Assignment delays,
- -Probability of arrival,
- -Probability of detection,
- -Probability of conversion,
- -Engagement type, range, aspect distribution,
- -Vectoring error.

## FUTURE UPGRADES

The basic capabilities at both AFEWES and REDCAP need major upgrades to represent potential weapon and C3 systems that our forces and aircraft will have to counter in the 1990s. Air Staff, SAF/AQQ, has directed that these upgrades be started immediately for delivery in the FY90-93 timeframe. For more information on the current and/or future capabilities of the AFEWES and REDCAP to support your test, contact AFOTEC/OL-AI, Lt Col Bryson, AV 838-5854.

**Lt. Col. Allen L. Bryson** is Director of the AFOTEC Operating Location AI, AF Plant #4, Ft. Worth, Texas. His responsibilities include planning for simulator upgrades and facilitating test conduct in support of operational testing at both the AFEWES and REDCAP facilities. His previous assignments have included duty in the B-52D (169 combat missions in SEA), RC-135V/M (117 operational missions), Flight Commander at the AF Electronic Warfare Training School, Senior Military ELINT Analyst for DIA Warnings and Indications, and ASD Program Manager for the AFEWES and REDCAP facilities. Lieutenant Colonel Bryson holds a bachelor's degree in Chemistry from Southern Nazarene University and a master's of System Management from the University of Southern California. He has been awarded two Distinguished Flying Crosses, the Defense Meritorious Service Medal, two Meritorious service Medals, and thirteen Air Medals.

# Relational Database Design

By Capt. Jeffrey A. Simmers,
Det 2 AFOTEC/ASPJ, Eglin AFB, FL

## INTRODUCTION

Databases are more than just organized data. A functional database contains both the plan for storing the data, typically called the database, and the application software that utilizes the data. This paper will call the application software database management software, regardless of whether it's user written or a commercial program. The primary topic of this paper is the design of relational databases, which are databases whose data structures follow the 'relational' model. Although this paper does not cover database management software in any detail, more information can be found in Reference 1.

## RELATIONAL DATABASES

Database architectures are usually classified into three major categories: relational, hierarchical (tree), and plex (network). Another category, called the inverted list, is actually an indexing scheme that can be used within any of the three and is not generally considered a separate architecture. This article addresses database development within the relational architecture, stressing the importance of normalization. (Ref 1)

One of the primary characteristics of a relational database is the storage of data in one or more flat files. A flat file stores data in two dimensional tables (rows and columns). The tables are structured so that the information about the relationships between data items is preserved, and the term 'relation' can be applied to the tables. These relations, or tables of data, have caused the database management programs that use such structures to be called relational database managers (rDBMS).

Another equally important concept behind relational databases involves the concept of data normalization. Normalization can be summarized as the building of the tables such that not only is all of the information regarding the relationships preserved, but that data redundancy is kept to a minimum, and the table rows and columns can be viewed in any sequence without affecting the functions that use the tables. Often literature explains normalization by constraining primary keys so that each different key value must uniquely identify one row of the table and each key cannot contain any redundant information. Implementing primary keys in this manner is the usual approach to achieving normalization, and will be demonstrated (along with the definition of what a primary key is) in the examples in the following section.

## DATA RELATIONS, NORMALIZATION AND KEYS

Although this article targets relational database management, many of the concepts and database building examples can be applied to tree and plex database architectures, because their structures can be represented as flat files. (Ref 1)

Consider the test team that wishes to build a database to store and manage the data represented in Table 1. The first question the team analyst asks is,

TABLE 1
Test Team Data

| Mission Number | Flown Date | System | ID |
|---|---|---|---|
| 100 | 01 JAN 87 | SADS I | .10 |
| 100 | 01 JAN 87 | SADS II | .20 |
| 100 | 01 JAN 87 | SADS III | .30 |
| 100 | 01 JAN 87 | SADS IV | .40 |
| 101 | 04 JAN 87 | SADS I | .11 |
| 101 | 04 JAN 87 | SADS II | .22 |
| 101 | 04 JAN 87 | SADS III | .33 |
| 101 | 04 JAN 87 | SADS IV | .44 |

will the team have the CRUD (create, retrieve, update, and delete) requirement for any or all of the data? If so, then building a normalized flat file database and using a relational database manager (such as DBASE III, CONDOR, or R:BASE V) should certainly be considered. Assuming the team decides to follow such a plan, let's step through the basic procedure to build such a database.

Table 1 shows a flat arrangement of the test team's effectiveness data into rows and columns.

Usually the rows are called records, and the columns are called fields. The database management software accesses the rows using record numbers, and the fields through field identifiers or titles. Table 1 shows the field titles, such as Mission Number, but does not show the record numbers, which are just sequential numbers given to each row starting with the first. Immediately we see that the table (or relation) contains two major types of data. The Mission Number and Flown Date fields are administrative, while the System and ID fields are effectiveness. Obviously, the combination of administrative and effectiveness data is redundant since the Mission Number and Flown Date are needlessly repeated. This relation is not considered "normalized" because of the redundancy.

The test team can eliminate this redundancy by separating the relation of Table 1 into two relations, shown in Tables 2 and 3. Table 2 contains the administrative relation, while Table 3 holds the effectiveness relation. Note that Table 2 still allows the test team to track the date that each mission was flown just as was originally done with the Flown Date field in Table 1.

TABLE 2

Administrative Relation

| Mission Number | Flown Date |
|---|---|
| 100 | 01 JAN 87 |
| 101 | 04 JAN 87 |

TABLE 3

Effectiveness Relation

| Mission Number | System | ID |
|---|---|---|
| 100 | SADS I | .10 |
| 100 | SADS II | .20 |
| 100 | SADS III | .30 |
| 100 | SADS IV | .40 |
| 101 | SADS I | .11 |
| 101 | SADS II | .22 |
| 101 | SADS III | .33 |
| 101 | SADS IV | .44 |

Using Table 2 we can now introduce the concept of a primary key. A primary key is one or more fields used to uniquely identify each record of the relation. The Mission Number is a logical primary key, since there is only one record for each Mission Number.

Choosing a primary key for Table 3 isn't so simple, since the Mission Number alone isn't enough to uniquely identify a record. However, by combining Mission Number and System we can create a primary key that does meet the uniquely identify requirement. Many database managers allow you to combine fields using string concatenation, and further permit you to convert a non-string field into a string field for concatenation. Combining Mission Number and System to create a primary key causes us to consider another important aspect of primary keys, that they should not contain redundant information. An example of redundancy within a key would be if we used Mission Number + Flown Date + System for a primary key in Table 1 (the "+" indicates concatenation). The Flown date would be redundant, since each mission number only has one possible Flown date. But Mission number + System in Table 3 does not contain redundant information, and so meets the requirements for a primary key.

By separating Table 1 into two relations, we have not only organized the data more logically, but have reduced our computer storage requirements (which may not always be the case). As previously discussed, we have also normalized the data relations. There are various degrees of normalization, but this article will just address the minimum degree required for most data base work. Further information on normalization can be found in References 1 and 2. Normalization is the organization of the relations so that each record can be uniquely identified by the primary key (no repeating records), and that every non-key field is fully functionally dependent upon the primary key (i.e. you can logically search for data in any field of the record by using the primary key). Both Tables 2 and 3 meet these two normalization requirements. Neither table repeats any records, and in both tables we could track down the information contained in the non-key fields by logical use of the primary key. For example, if we wanted to know what the ID values were against the SADS II for all missions, we could easily obtain this information using just the primary key.

Another type of key, called the secondary key, is sometimes used to allow the user to extract certain types of information from a relation, or

to logically tie two different relations together. In Table 3 a candidate secondary key would be the Mission Number, as each Mission Number in this table could be equated with a Mission number in Table 2. Usually the secondary key does not meet all of the requirements for a primary key. Not only does Mission Number allow us to tie the two tables together (effectively creating a larger 'logical' table), but it also lets us extract information from Table 3 for each mission. A secondary key *used to tap* into another relation may also be called a foreign key.

## PERFORMANCE

After the user initially structures the database based upon data relations, normalization, and primary/secondary keys, this structure must be revisited based upon several other important considerations. The first of these is database performance, usually in terms of optimizing the speed of as many of the planned database tasks as possible. Database tasks are basically simple or complex combinations of the CRUD operations. A simple task would be the addition of new records into the database, while a more complex task would be producing averages of multiple fields within multiple records. General guidelines for improving database performance are:

-With either real or created data, build the initial data base structure.
-Time as many of the planned tasks as feasible using this initial database.
-Extrapolate the timing data to account for the projected size of the actual database.
-Decide which tasks should be optimized.
-Redesign the database structure to increase the speed of the tasks marked for optimization.

These five steps can be repeated over and over until the user is satisfied that no further progress can be made, i.e., this is a 'recursive' process.

As an example, consider the relation in Table 4. The primary task performed by the test team on the database will be to calculate the total amount of money spent on all items. Using the structure of Table 4, the team will have their database management software multiply the cost times the quantity for each item, and then add up all the subtotals. Typically, the team enters one new

**TABLE 4**
**Cost Data**

| Item | Cost | Quantity |
|---|---|---|
| 1A2 | 5.00 | 100 |
| 2A1 | 7.00 | 100 |
| 3A6 | 8.00 | 100 |

item per day, and calculates the total once at the end of each month. The result of their timing runs on a small test data base yields the following data:

| TASK | TIME |
|---|---|
| Enter one item | 1 minute |
| Total ten items | 10 minutes |

Unfortunately, the team projects that by the end of their test they will have at least 1000 items in their database. *Extrapolating the time* to total the cost for ten items to that for 1000 items yields 16 hours and 40 minutes! Obviously this task is a candidate for performance optimization through database restructure. Since computing the total cost involves the calculation of cost per item times number of items for each record, the team decides to create a new field for each record to store this quantity. They modify their database management software to automatically calculate this value and enter it into the database once the user has keyed in the item and its price. Table 5 shows the new database with this new field called Subtotal, which is a derived field, because it is derived from other fields instead of keyed in directly by

**TABLE 5**
**Revised Cost Data**

| Item | Cost | Quantity | Subtotal |
|---|---|---|---|
| 1A2 | 5.00 | 100 | 500.00 |
| 2A1 | 7.00 | 100 | 700.00 |
| 3A6 | 8.00 | 100 | 800.00 |

the user. New timing runs yield the following results:

| TASK | TIME |
|---|---|
| Enter one item | 1.5 minutes |
| Total ten items | 5 minutes |

Now the projected time to calculate the end of month cost totals for a 1000 item database is cut in half to eight hours and 20 minutes.

Although this example may seem almost too simple to take seriously, in practice the performance optimization steps can be quite sticky. Typically users will build a database before they completely define the tasks to be performed, complicating step two. Step 3 requires some analysis to decide upon the shape of the extrapolations (i.e. linear versus non linear increase in time versus numbers of records). Additionally, step four forces the user to prioritize all of the planned tasks, and consider how and when they will all be accomplished.

## SECURITY

Security includes two major categories: security against unauthorized access, and data integrity. Both categories generate two questions that need to be answered: What should be protected; how should it be protected? As part of the answer to the second question the user should consider alternative methods of protection, and the relative costs (time, money, effort) of each.

Consider again the database from Table 5. Our test team has optimized the database structure for performance, but a security problem centers around the Cost field. Since these items are part of a competitive test for possible procurement by the Air Force, the team must restrict the ability to extract the cost of each item to just test team members. One way to do this would be to use password protection in the database management software that allows viewing of the Cost and Subtotal fields, while allowing free access to the Item and Quantity fields.

Another aspect of security is data integrity. Returning to the database of Table 5, our team discovers that an erroneous entry for the Cost field could result in a substantial error in the monthly cost total. The lack of built in safeguards for data integrity is a major shortcoming of many locally (i.e. by a test team or office) designed databases.

For our test team, one solution to these security concerns is a restructuring of the single data base of Table 5 into two separate relations. One relation contains only the Item and Quantity data, while the other is identical to the relation of Table 5 but with the addition of two software security safeguards. The first safeguard is password protection for any use of the database, while the second is an 'edit check' on the Cost field. This edit check would check every value entered into the Cost field against a rule, such as:

$$\text{Cost} <= 10.00$$

Edit checks, like this one against the Cost field, come from a prior knowledge of the values for a given field, such as our team's knowing that no single item can cost more than ten dollars. This edit check would catch some errors, but obviously will not catch an erroneous item cost of less than ten dollars.

The splitting up of the database from Table 5 does not have to be physical. The team could still have one physical database, but modify their database management software so that the users of the database would think there were actually two databases, with the primary difference that one database would require a password and the other wouldn't. If, however, the team elected to keep two physically separate databases, then another aspect of data integrity must be considered. If the smaller database (Item and Quantity fields only) can be changed by a user, then the team could easily have the situation where the two databases don't have the same data in their common fields. To prevent this the team can only allow the smaller database to be read from, not written to, by the user. Users with the correct password could update the full database (the one with the costs), and then the database management software would automatically update the smaller database.

One final caution about data integrity. With the proliferation of math co-processor chips, the probability of the computer actually making an arithmetic mistake seems to have risen somewhat. This is due to the independence of operation between the main processor chip and the math co-processor chip, allowing the main processor to be unaware of a problem with the math chip.

One possible guard against arithmetic error would be a technique devised for the first electronic digital computer built in this country, the ENIAC. ENIAC programmers routinely ran an arithmetic intensive program (with a known solution) through their machine as a check for faulty operation (Ref 3). Our test team could easily do the same by periodically running the cost totaling program on a non-changing baseline data base to check for numerical accuracy.

## GROWTH

Failure to adequately project the growth of a database can lead to major complications as the database grows beyond the processing or storage capabilities of the computer and/or the database management software. Growth projections used for performance analysis can yield valuable information pertaining to the capability of the present hardware and software to handle the future database. An office or test team preparing a request for computer equipment should consider if the requested equipment will be capable of meeting projected database *management requirements*.

If database growth appears destined to outstrip the local computer resources, then consider alternatives, such as the use of a time-shared mainframe computer. One alternative would be storing raw data on a mainframe computer, then periodically downloading either processed data or subsets of the raw data for management on local PCs (personal computers). Integrating different databases in different locations in this manner is called "distributed" database processing, which is a complex topic in itself (Ref 1).

Another approach to compensate for problems due to database growth is manipulating the database structure. Breaking up one large relation into smaller relations, as we did with Table 1, will often solve the problem. For a database with substantial derived data, separating the derived fields into a separate relation may help. Unfortunately, modifications made to facilitate growth considerations may adversely impact previous changes made for other reasons.

## SUMMARY

This paper presents guidelines for structuring the data within the relational database model to facilitate the concepts of normalization, keys, performance, security, and growth. But database structure is just the first of a two part process. The second step is the design and implementation of the database applications using database management software. This software may be a commercial database manager, such as dBASE III, or it may be a specialized program written by the user. Although this paper does not address such software applications, database structure and application software cannot be considered independently. Before designing a database structure, the user should consider the applications and their impact on the database design. Reference 2 contains more information about the integration of data and application.

## REFERENCES

1. Martin, James. Computer Data-Base Organization. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1977.

2. Seminar Topic: Data Base: A Builder's Guide. Delivered by Robert Holland and William Smith on a periodic basis, sponsored by Technology Transfer Institute, 741 Tenth Street, Santa Monica, CA 90402. Copyright 1986 by Robert Holland and William Smith.

3. Brainerd, J. G. and Sharpless, T. K. "The ENIAC," Proceedings of the IEEE, Vol. 72, No. 9 (Sept. 1984), 1203-1212.

**Jeffrey A. Simmers** is the Airborne Self protection Jammer test team Assistant for Analysis. He came to AFOTEC from the Air Force Institute of Technology (AFIT) at Wright-Patterson AFB, Ohio and before that was an Avionic Systems Engineer in the Avionics branch at the Foreign Technology Division (FTD), also at Wright-Patterson. Captain Simmers was enrolled in the Guidance and Control sequence while attending AFIT and wrote his thesis on the Flight Propulsion Control Coupled (FPCC) aircraft. At FTD he analyzed foreign communication and navigation systems and wrote a Defense Intelligence Agency study plus several foreign material exploitation reports. Captain Simmers holds a bachelor's in Electrical Engineering from the University of Illinois, a master's in Electrical Engineering from AFIT and will be entering a Ph.D. program this July.